

Intro to data analysis with R.

Ronny A. Hernández Mora.

@RonnyHdezM

ronnyhdez

<http://ronnyhdez.rbind.io/>

Workshop materials

README.md

¡Bienvenido al curso de ciencia de datos con R!

Este es un curso libre y gratis que puede usar para dar sus primeros pasos con el lenguaje de programación R.

A cartoon illustration of a hot air balloon with a speech bubble from a character on a trampoline. The hot air balloon has two baskets, each containing a character. One basket has a speech bubble that says "I'm doing a thing all on my own!". The other basket has a speech bubble that says "code hero". Below the hot air balloon is a blue trampoline with three stick figures. One figure is jumping on the trampoline, another is standing next to it, and a third is sitting on the left side. The trampoline has a speech bubble that says "support" and another that says "community". The artist's name, @allison_horst, is at the bottom right of the illustration.

Artwork by @allison_horst

Materiales del curso

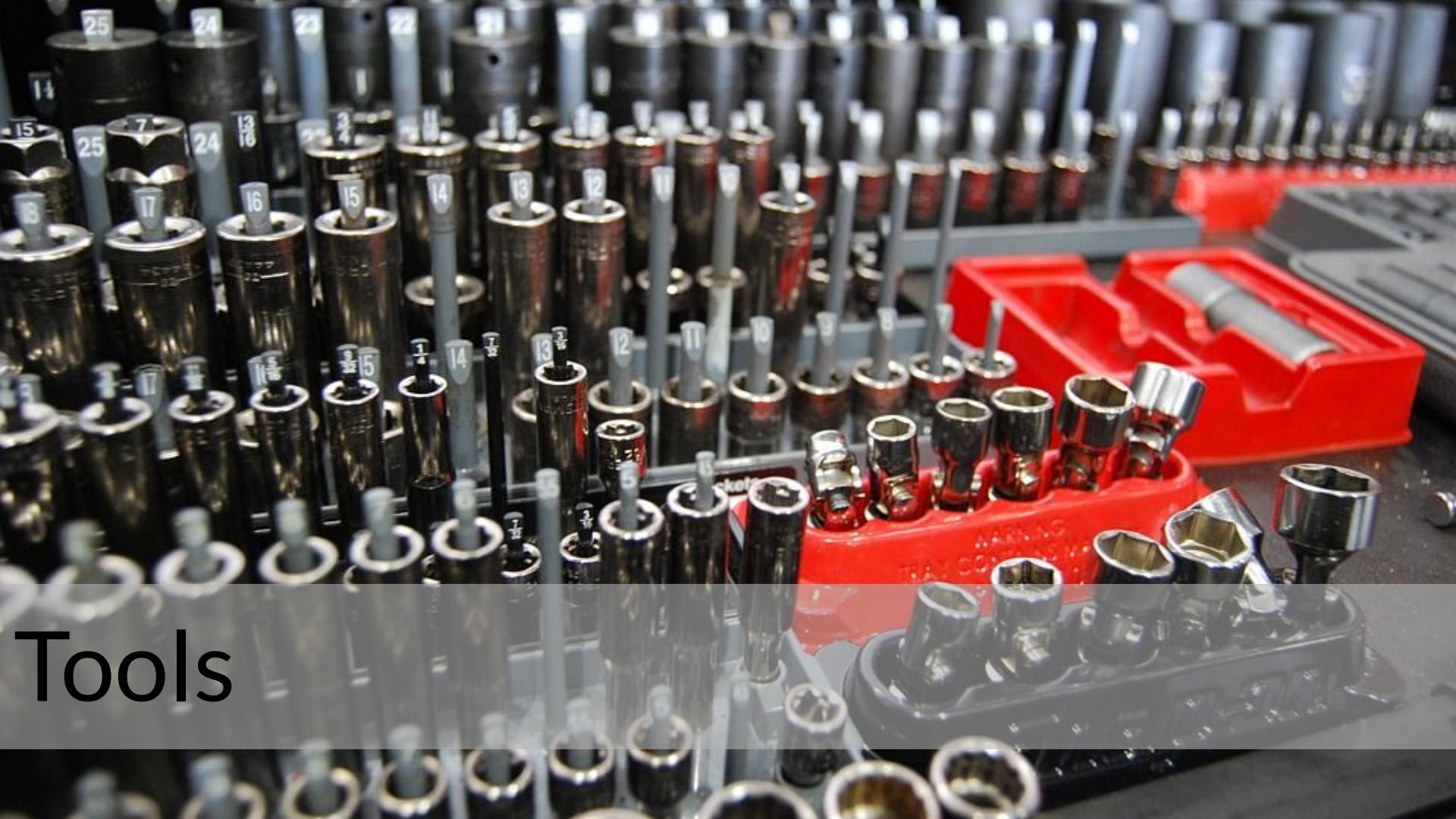
Sesión	Presentación	Video
0-Preparación	No hay	https://www.youtube.com/watch?v=NCvJwJSMq60
1- Introducción a las herramientas	Por subir	Por subir



https://github.com/ronnyhdez/intro_data_analysis

What do we want from today's session?

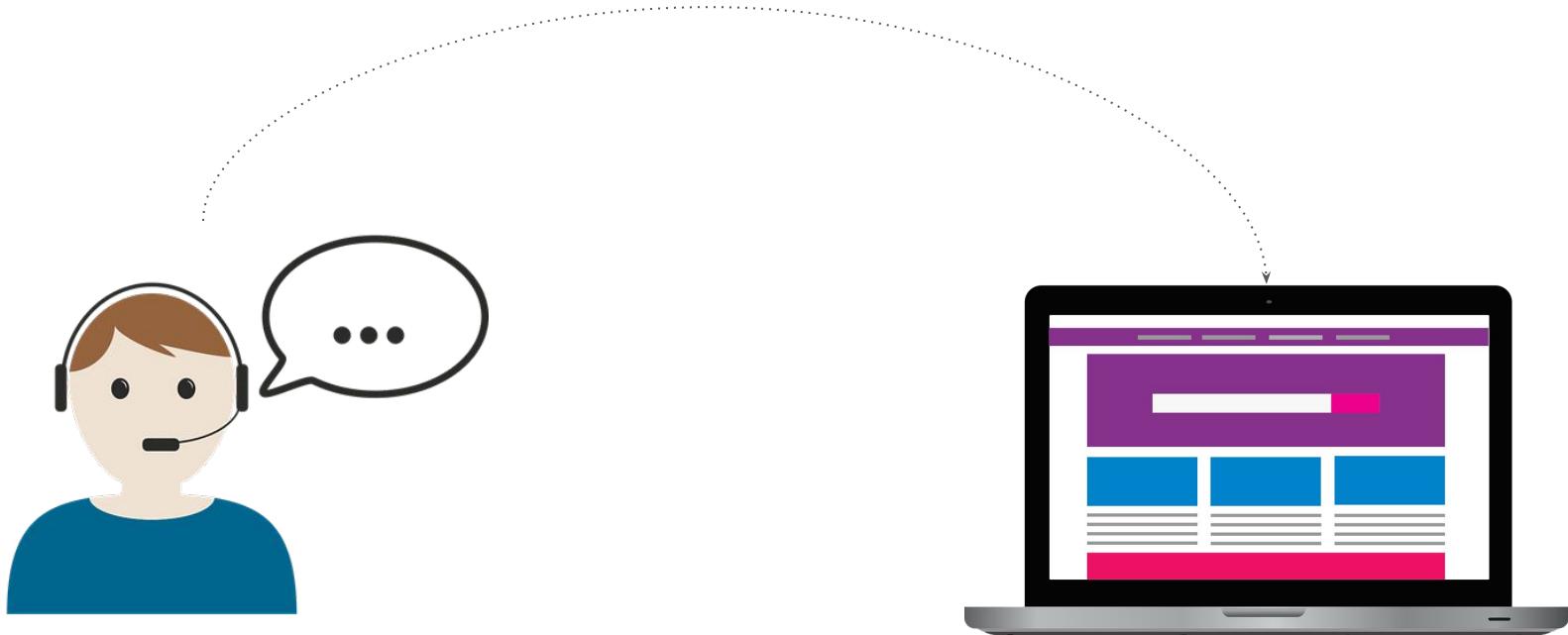
- To understand what is R and why is useful
- To understand the relationship between R and RStudio
- Start an R project
- Navigate in RStudio
- Understand R objects and environment

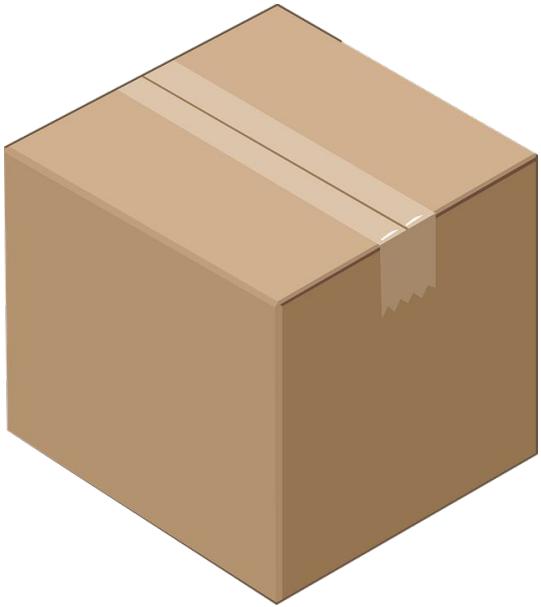


Tools

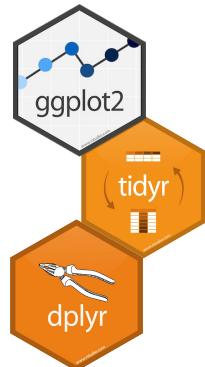
Definition from [R Project for Statistical Computing](#):

R is a programming language and free software environment for statistical computing and graphics.



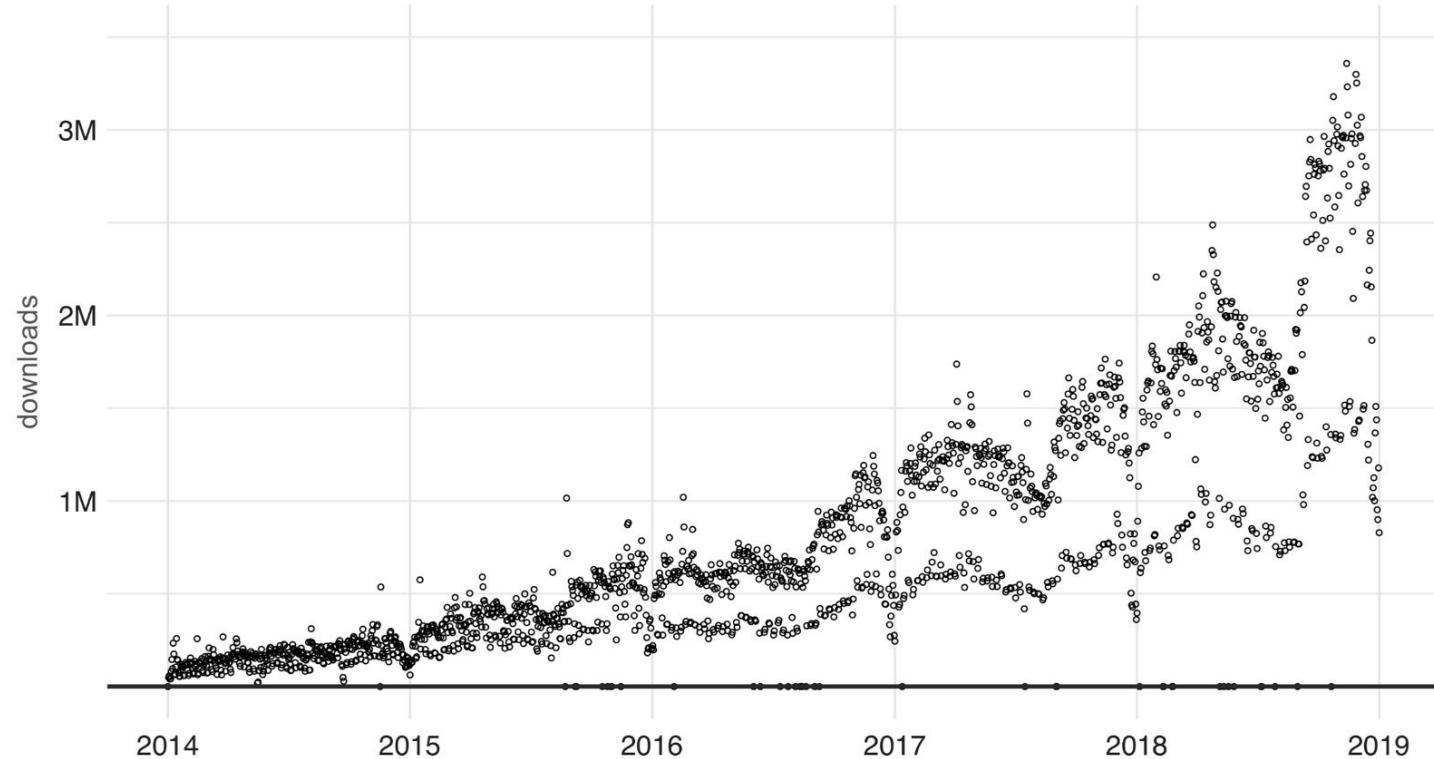


Packages in R



CRAN Packages

Total daily downloads over time





[Home]

[Download](#)

[CRAN](#)

[R Project](#)

[About R](#)

[Logo](#)

[Contributors](#)

[What's New?](#)

[Reporting Bugs](#)

[Conferences](#)

[Search](#)

[Get Involved: Mailing Lists](#)

[Developer Pages](#)

[R Blog](#)

[R Foundation](#)

[Foundation](#)

[Board](#)

[Members](#)

[Donors](#)

[Donate](#)

[Help With R](#)

[Getting Help](#)

[Documentation](#)

[Manuals](#)

[FAQs](#)

[The R Journal](#)

[Books](#)

[Certification](#)

[Other](#)

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

News

- R version 4.1.0 (Camp Pontanezen) has been released on 2021-05-18.
- R version 4.0.5 (Shake and Throw) was released on 2021-03-31.
- Thanks to the organisers of useR! 2020 for a successful online conference. Recorded tutorials and talks from the conference are available on the R Consortium YouTube channel.
- You can support the R Foundation with a renewable subscription as a supporting member.

News via Twitter

The R Foundation Retweeted

Peter Dalgaard @pdalgd #rstats 4.1.0 "Camp Pontanezen" (source version) has been released.

[1] 3

16h

The R Foundation Retweeted userR! 2021 @_useRconf After a 24h maintenance break, registrations are open again!

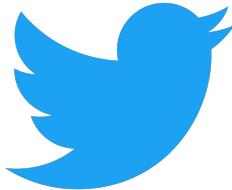
Official packages that goes through a review process



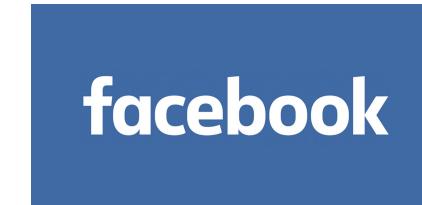
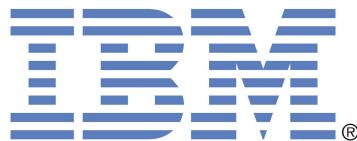
others...

Packages being develop

¿Quienes usan R?



THE WORLD BANK



Who is using R?

National Ecological Observatory Network

Boulder, CO <http://www.neonscience.org/>

Overview Repositories 49 Packages People Projects

Pinned

NEON-utilities Public

Utilities and scripts for working with NEON data. Currently: an R package with functions to join (stack) the month-by-site files in downloaded NEON data, to convert data to geoCSV format, and to do...

R 46 29

Repositories

Find a repository... Type Language Sort

DI-NEON-participants

A repository used for training and collaboration in NEON Data Institutes and other workshops.

HTML 8 138 1 (1 issue needs help) 9 Updated 31 minutes ago

NEON-stream-discharge

Code and packages for NEON stream discharge calculations. Includes (1) the code package used to generate NEON's Stream discharge rating curve (DP4.00133.001) and Stream discharge (DP4.00130.001) data products and (2) a package for calculating point estimates of discharge from salt injections.

R 5 APL3.0 4 0 0 Updated 2 days ago

neon-plant-sampling

Scripts supporting NEON plant biomass and productivity sampling

People

This organization has no public members. You must be a member to see who's a part of this organization.

Top languages

R HTML JavaScript Python Jupyter Notebook

Most used topics

ecology neon

Report abuse

<https://github.com/neonscience>

What can we do with R?

Nutrition Calculator

1. Ingredient

Type to search for ingredient

Add ingredient

Remove ingredient

Number of servings contained

1

Note: All nutrient information is based on the Canadian Nutrient File. Nutrient amounts do not account for variation in nutrient retention and yield losses of ingredients during preparation. % daily values (DV) are taken from the Table of Daily Values from the Government of Canada. This data should not be used for nutritional labeling.

kcal

Calories

All nutrients below recommended DV

All nutrients below 50% recommended DV

Ingredients

Show 5 entries Search:

quantity units ingredient_name

No data available in table

Showing 0 to 0 of 0 entries Previous Next

Macronutrients

% Daily Value

Cholesterol Fat Fibre Sodium Sugars Saturated and Trans Fats Nutrient

Nutrition Table

for 1 serving (1 servings in recipe)

Show 10 entries Search:

NutrientName Value Unit

No data available in table

Showing 0 to 0 of 0 entries Previous Next

Minerals

Value

<https://shiny.rstudio.com/gallery/nutrition-calculator.html>

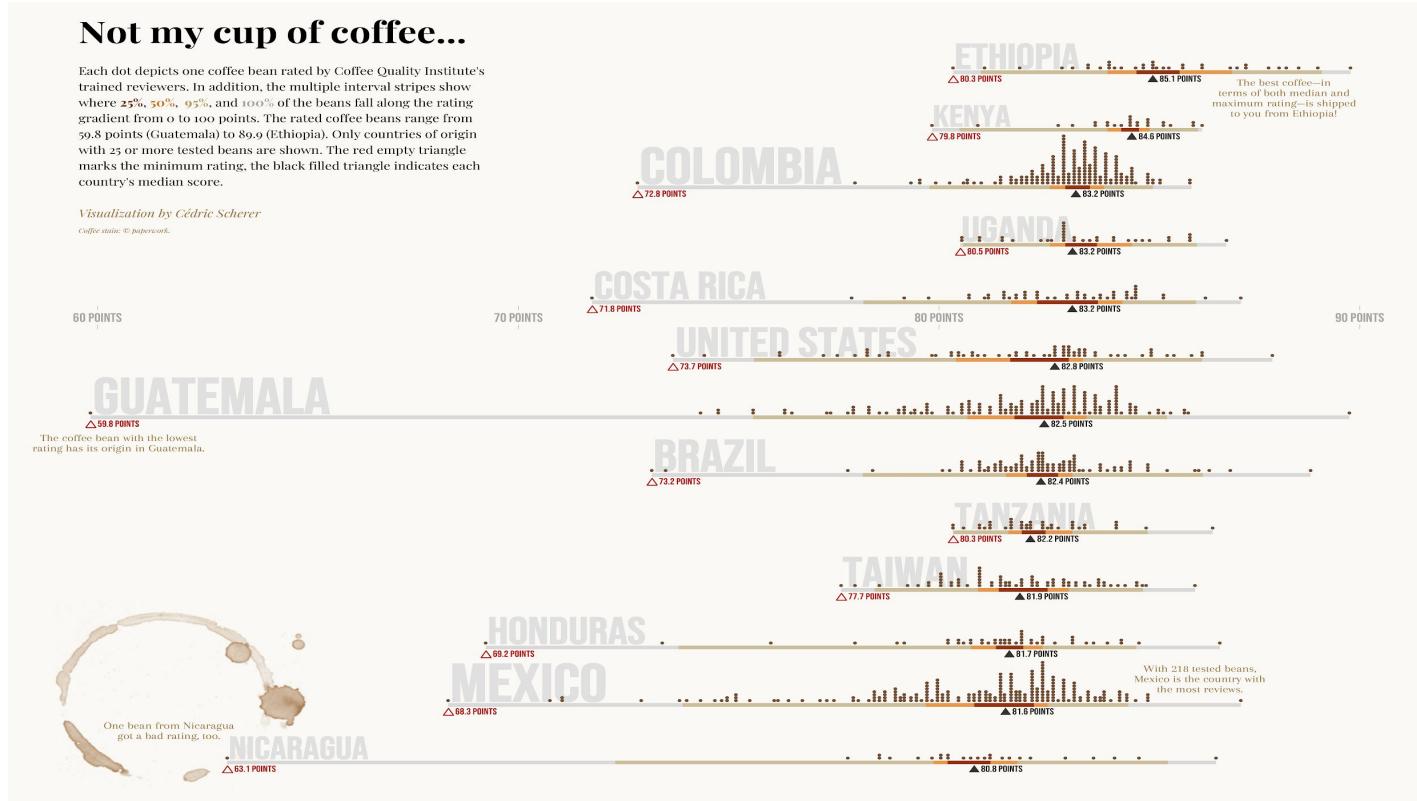
What can we do with R?

Not my cup of coffee...

Each dot depicts one coffee bean rated by Coffee Quality Institute's trained reviewers. In addition, the multiple interval stripes show where 25%, 50%, 75%, and 100% of the beans fall along the rating gradient from 0 to 100 points. The rated coffee beans range from 59.8 points (Guatemala) to 89.9 (Ethiopia). Only countries of origin with 25 or more tested beans are shown. The red empty triangle marks the minimum rating, the black filled triangle indicates each country's median score.

Visualization by Cédric Scherer

Coffee static: © paperwork.



¿Qué se puede crear con R?

1 Consideraciones:
2 Date ranges data visualization
3 Evapotranspiration calculation

Code ▾

Evapotranspiration calculation

EAS 520

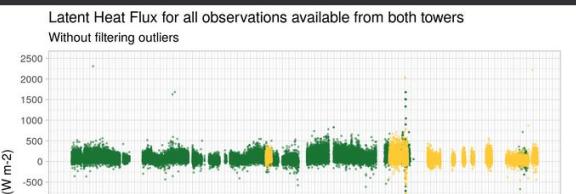
Ronny Hernández Mora
2/27/2021

1 Consideraciones:

- Para los datos de temperatura, si tenemos muchos valores por minuto, por hora por día, agruparlos y sacar un promedio.
- Luego, estos valores promediados por minutos, promediárlas por hora o por día para hacer el cálculo.
- Tenemos que unir los conjuntos de datos que contienen latent_heat_flux y temperatura para conocer cuál es el período de tiempo más largo posible que tenemos.
- Con evapotranspiración calculada, sacar conclusiones

2 Date ranges data visualization

Latent Heat Flux for all observations available from both towers
Without filtering outliers



https://github.com/ronnyhdez/eas_520

What other technologies can be integrated with R?



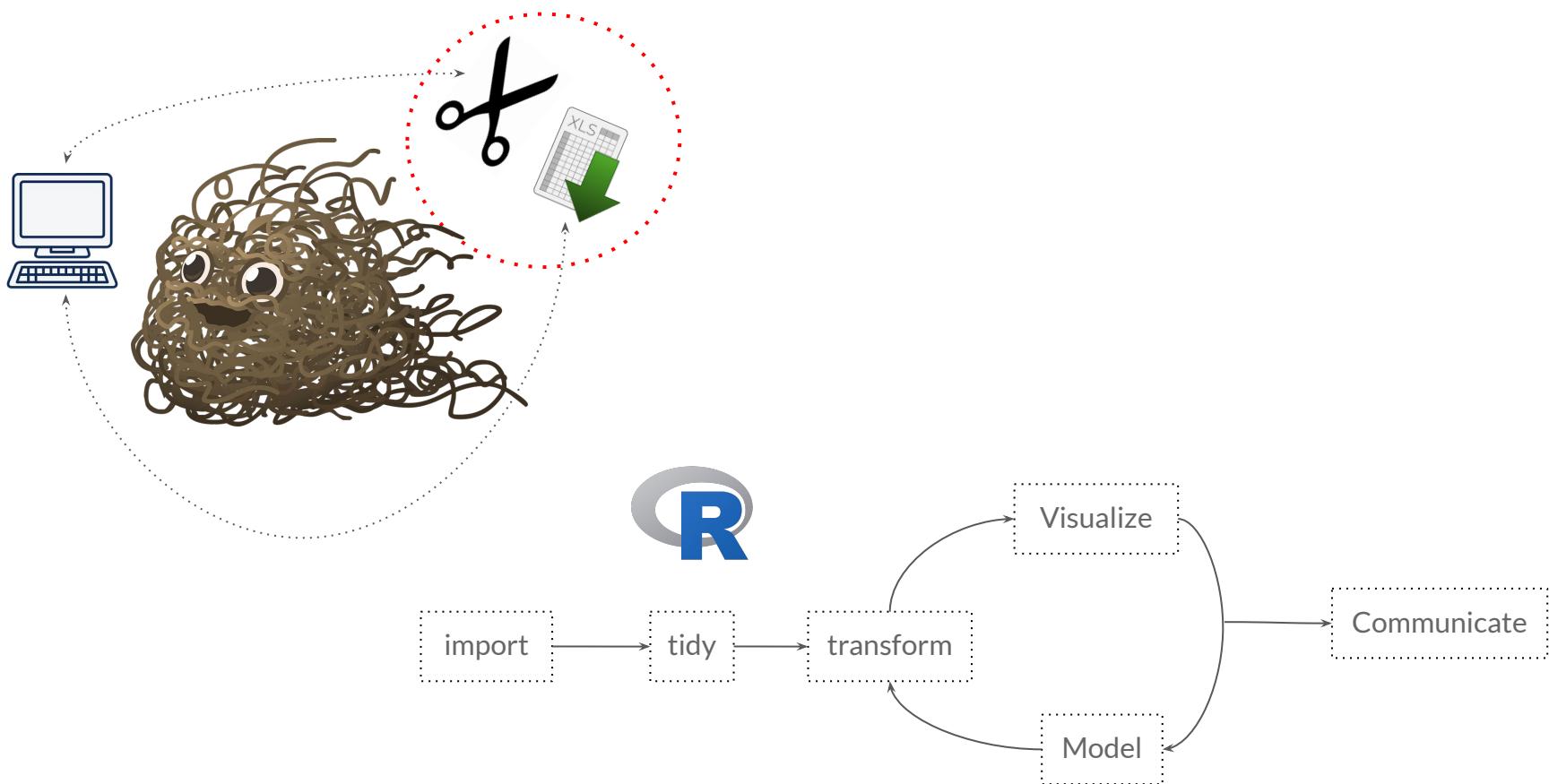
Pachyderm
version-controlled data science



kubernetes

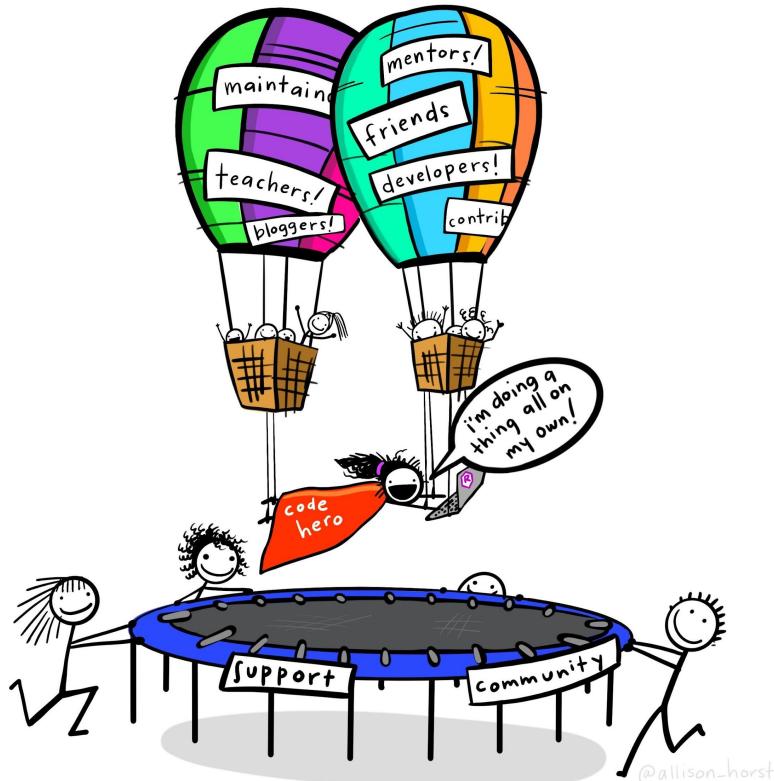


Advantages on using R instead of GUI's





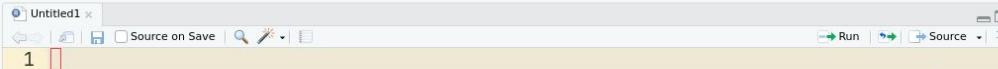
Other advantages of using R



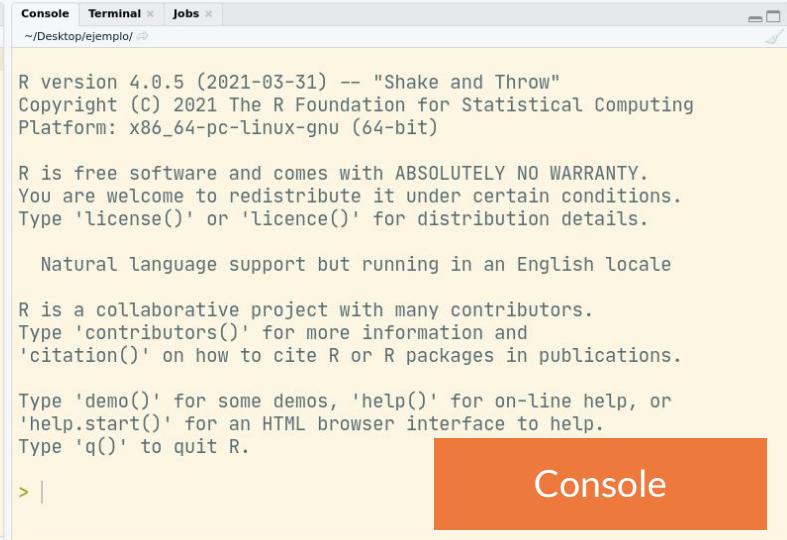
Artwork by @allison_horst



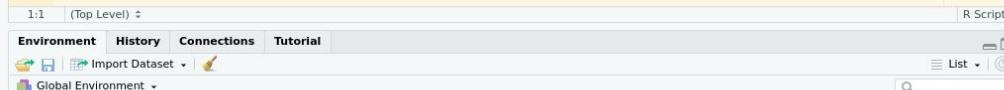
Its an Integrated Development Environment (IDE) that allows us to have an easier workflow.



Editor



Console



Environment



Output



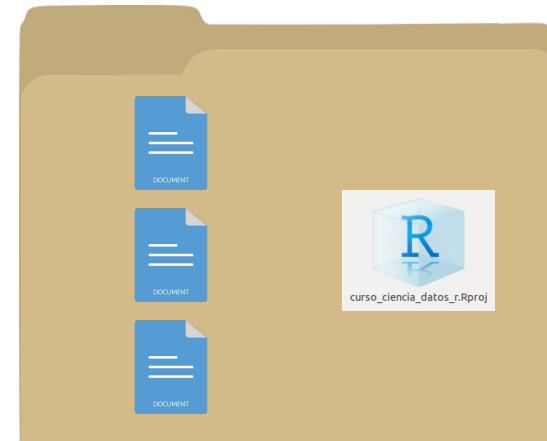
Break 12:12 a 12:22

Intro to R



Artwork by @allison_horst

Primero: flujo de trabajo



```
my_object <- "hola mundo"
```

```
Object_3-WearName <- "adios mundo"
```

```
# This is a comment
```

```
3 * 4
```

```
my_result <- 3 * 4
```

```
seq(1, 10)
```

```
seq(1, 10, by = 2)
```

```
?seq()
```

Function name

```
seq( from = 1 , to = 10 )
```

Argument 1 Argument 2

Separate arguments

Function parenthesis

```
seq(1, 10)
```

```
seq(from = 1, to = 10)
```

```
seq(1, 10, by = 2)
```

```
?seq()
```



Supplementary text

The R Series

R Markdown

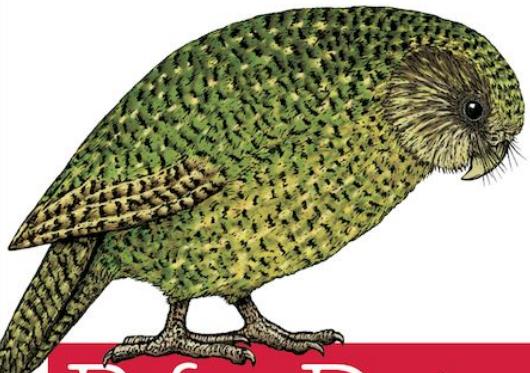
The Definitive Guide



Yihui Xie
J. J. Allaire
Garrett Grolemund

 CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

O'REILLY®



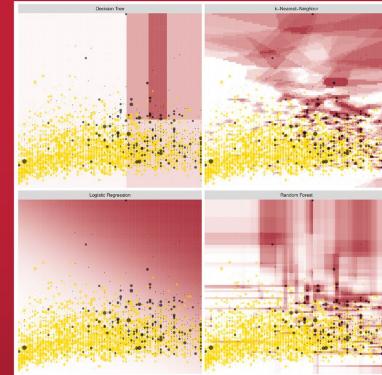
Hadley Wickham &
Garrett Grolemund

<https://es.r4ds.hadley.nz/>

Texts in Statistical Science

Modern Data Science with R

Second Edition



Benjamin S. Baumer
Daniel T. Kaplan
Nicholas J. Horton

 CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

<https://es.r4ds.hadley.nz/>

<https://bookdown.org/yihui/rmarkdown/>



Thanks!

Ronny A. Hernández Mora.



@RonnyHdezM

ronnyhdez

<http://ronnyhdez.rbind.io/>